A Vector-Valued Almost Sure Invariance Principle for Hyperbolic Dynamical Systems

Ian Melbourne * Matthew Nicol †

20 June, 2006

Abstract

We prove an almost sure invariance principle (approximation by d-dimensional Brownian motion) for vector-valued Hölder observables of large classes of nonuniformly hyperbolic dynamical systems. These systems include Axiom A diffeomorphisms and flows as well as systems modelled by Young towers with moderate tail decay rates.

In particular, the position variable of the planar periodic Lorentz gas with finite horizon approximates a 2-dimensional Brownian motion.

1 Introduction

The scalar almost sure invariance principle (ASIP), or approximation by one-dimensional Brownian motion, is a strong statistical property of sequences of random variables introduced by Strassen [40, 41]. It implies numerous other statistical limit laws including the central limit theorem, the functional central limit theorem, and the law of the iterated logarithm. See [23, 38] and references therein for a survey of consequences of the ASIP.

The scalar ASIP has been shown to hold for large classes of dynamical systems [14, 17, 18, 21, 24, 25, 30, 31, 35]. Chernov & Dolgopyat [10, Problem 1] asked for a proof of the ASIP for $\mathbb{R}^d$-valued observables, and it is this problem that is solved in this paper. Our main result applies to a large variety of dynamical systems, as surveyed in Section 4.

As a secondary matter, we obtain explicit error estimates that depend on the dimension d and the lack of hyperbolicity. Even for $d = 1$, this estimate is better than those in almost all of the above references. The exception is [21] which gives the best available estimate for scalar ASIPs for a restricted class of systems.

*Department of Mathematics and Statistics, University of Surrey, Guildford GU2 7XH, UK. E-mail: ism@math.uh.edu

†Department of Mathematics, University of Houston, Houston TX 77204-3008, USA. Email: nicol@math.uh.edu

1.1 Statement of the main results

Definition 1.1 A sequence $\{S_N\}$ of random variables with values in $\mathbb{R}^d$ satisfies a *d-dimensional almost sure invariance principle (ASIP)* if there exists $\lambda > 0$ and a probability space supporting a sequence of random variables S_N^* and a d-dimensional Brownian motion $W(t)$ such that

(a) $\{S_N; N \geq 1\} =_d \{S_N^*; N \geq 1\}$, and

(b) $S_N^* = W(N) + O(N^{\frac{1}{2}-\lambda})$ as $N \to \infty$ almost everywhere.

For brevity, we write $S_N = W(N) + O(N^{\frac{1}{2}-\lambda})$ a.e. The ASIP for a one-parameter family S_T of $\mathbb{R}^d$-valued random variables is defined similarly, and denoted $S_T = W(T) + O(T^{\frac{1}{2}-\lambda})$ a.e.

Remark 1.2 The ASIP is said to be *nondegenerate* if the Brownian motion $W(t)$ has nonsingular covariance matrix Σ. For the classes of dynamical systems considered in this paper, the ASIP is nondegenerate for *typical* observables. More precisely, there is a closed subspace Z of infinite codimension in the space of all (piecewise) Hölder $\mathbb{R}^d$-valued observables such that Σ is nonsingular whenever $\phi \notin Z$. (By considering all one-dimensional projections it suffices to consider the case $d = 1$. This is done explicitly in for example [25, Section 4.3].)

Axiom A diffeomorphisms and flows Our results are most easily stated in the uniformly hyperbolic (Axiom A) context.

Theorem 1.3 *Let $f : M \to M$ be a diffeomorphism with a (nontrivial) uniformly hyperbolic basic set $X \subset M$, and suppose that μ is an equilibrium measure corresponding to a Hölder potential. Let $\phi : X \to \mathbb{R}^d$ be a mean zero Hölder observable with partial sums $S_N = \sum_{n=1}^N \phi \circ f^j$. Then for any $\epsilon > 0$,*

$$S_N = W(N) + O(N^{\beta+\epsilon}) \text{ a.e.} \quad \text{where } \beta = \tfrac{2d+3}{4d+7}.$$

(The improved estimate $\beta = \frac{1}{4}$ holds when $d = 1$ [21].)

An immediate consequence (see for example [17, 32]) is the corresponding result for Axiom A flows.

Corollary 1.4 *Let $f_t : M \to M$ be a smooth flow with a (nontrivial) uniformly hyperbolic basic set $X \subset M$, and suppose that μ is an equilibrium measure corresponding to a Hölder potential. Let $\phi : X \to \mathbb{R}^d$ be a mean zero Hölder observable with partial sums $S_T = \int_0^T \phi \circ f_t \, dt$. Then for any $\epsilon > 0$,*

$$S_T = W(T) + O(T^{\beta+\epsilon}) \text{ a.e.} \quad \text{where } \beta = \tfrac{2d+3}{4d+7}.$$

(The improved estimate $\beta = \frac{1}{4}$ holds when $d = 1$ [21, 32].) ∎

Remark 1.5 Denker & Philipp [17] proved Theorem 1.3 and Corollary 1.4 in the case $d = 1$ (though with a weaker error term).

Nonuniformly hyperbolic systems Our results apply also to maps $f : M \to M$ that are nonuniformly expanding/hyperbolic in the sense of Young [44, 45]. Roughly speaking, such maps possess a subset $\Lambda \subset M$ and a return time $R : \Lambda \to \mathbb{Z}^+$ such that the induced map $f^R : \Lambda \to \Lambda$ is uniformly hyperbolic.

Theorem 1.6 *Let $f : M \to M$ be a diffeomorphism (possibly with singularities) that is nonuniformly hyperbolic in the sense of Young [44, 45]. In particular, f satisfies conditions (A1)–(A4) in Section 4.2 and possesses an SRB measure m. Assume that the return time function R lies in L^p, $p > 2$. Let $\phi : M \to \mathbb{R}^d$ be a mean zero Hölder observation with partial sums $S_N = \sum_{n=1}^N \phi \circ f^j$. Then for any $\epsilon > 0$,*

$$S_N = W(N) + O(N^{\beta+\epsilon}) \;\; a.e. \quad where \;\; \beta = \frac{\frac{1}{p}+2d+3}{4d+7}.$$

If $d = 1$ then β can be improved to $\beta = \frac{1}{2p} + \frac{1}{4}$ for $2 < p \le 4$ and $\beta = \frac{3}{8}$ for $p \ge 4$.

Again, there is an immediate extension to nonuniformly hyperbolic flows. Suppose that $f : M \to M$ satisfies the assumptions of Theorem 1.6 with $R \in L^p$, $p > 2$, and that f_t is a suspension flow over f with a (uniformly bounded) Hölder roof function. By [32], $\mathbb{R}^d$-valued Hölder observables of the suspension flow satisfy an ASIP of the form $S_T = W(T) + O(T^{\beta+\epsilon})$ a.e. where β is as in Theorem 1.6.

Application to Lorentz gases The planar periodic Lorentz gas was introduced by Sinai [39]. This is a three-dimensional flow with phase space $(\mathbb{R}^2 - \Omega) \times S^1$, where $\Omega \subset \mathbb{R}^2$ is a periodic array of disjoint convex regions with C^3 boundaries. The coordinates are position $q \in \mathbb{R}^2 - \Omega$ and velocity $v \in S^1$. The flow satisfies the finite horizon condition if the time between collisions with $\partial\Omega$ is uniformly bounded.

Let $q(t) \in \mathbb{R}^2$ denote the position at time t of a particle starting at position $q(0)$ pointing in direction $v(0)$. Bunimovich & Sinai [6], see also [7], proved that $q(t)$ satisfies a two-dimensional functional central limit theorem (weak invariance principle) supporting the view of such flows as a deterministic model for Brownian motion. We complete this circle of ideas by proving the strong version of this result.

Theorem 1.7 *Consider a planar periodic Lorentz gas satisfying the finite horizon condition. Let $\epsilon > 0$. There is a two-dimensional Brownian motion $W(t)$ with nonsingular covariance matrix such that for almost every initial condition, $q(T) = W(T) + O(T^{\frac{7}{15}+\epsilon})$.*

Remark 1.8 (a) A number of authors [9, 30, 35] have independently established scalar ASIPs for one-dimensional projections of $q(t)$. In hindsight, the scalar ASIP for the Lorentz

gas follows from earlier work of [21], again with $\beta = \frac{1}{4}$. (We note that the methods of [21] apply in the first place only to the time-reversal of the dynamical system under study. Their applicability here is due to the fact that the class of systems is closed under time-reversal.)

(b) The finite horizon condition is crucial. For infinite horizons, Szász & Varjú [42] prove that $q(t)$ lies in the nonstandard domain of the normal distribution. In particular, the central limit theorem fails, hence the ASIP fails.

1.2 Consequences of the vector-valued ASIP

For convenience, we suppose that the Brownian motion in the ASIP is nondegenerate. Coordinates can be chosen on $\mathbb{R}^d$ so that $W(t)$ is a standard d-dimensional Brownian motion with $\Sigma = I_d$. Throughout, the norm on $\mathbb{R}^d$ is taken to be the usual Euclidean norm. The following consequences of the ASIP are summarised in [34, p. 233]. Here, LIL stands for *law of the iterated logarithm* and the functional LIL stated below is a far-reaching generalisation, due to Strassen, of the classical LIL.

Proposition 1.9 *For the dynamical systems to which the results in this paper apply, the following consequences hold (after normalisation so that $\Sigma = I_d$):*

• **Functional LIL** *Let $C = C([0,1], \mathbb{R}^d)$ be the Banach space of continuous maps $f : [0,1] \to \mathbb{R}^d$ with the supremum norm. Let K be the (compact) set of $f \in C$ absolutely continuous with $f(0) = 0$, $\int_0^1 |f'(t)|^2 dt \le 1$. Define $f_n(i/n) = S_i/\sqrt{2n \log\log n}$, $i = 0, \ldots, n$, and linearly interpolate to obtain $f_n \in C$. Then a.s. the sequence $\{f_n\}$ is relatively compact in C and its set of limit points is precisely K.*

• **Upper and lower class refinement of the LIL** *Let $\phi(t) : \mathbb{R} \to \mathbb{R}$ be a positive nondecreasing function. Then*

$$P(|S_N| > N^{\frac{1}{2}}\phi(N) \ \ i.o.) = 0 \ \ or \ \ 1$$

according to whether $\int_1^\infty \frac{\phi^d(u)}{u} \exp(-\frac{1}{2}\phi^2(u))\, du$ converges or diverges.

• **Upper and lower class refinement of Chung's LIL** *Let $\phi(t) : \mathbb{R} \to \mathbb{R}$ be a positive nondecreasing function. Then there is a constant c (depending only on d) such that*

$$P(\max_{n \le N} |S_n| < cN^{\frac{1}{2}}\phi^{-1}(N) \ \ i.o.) = 0 \ \ or \ \ 1$$

according to whether $\int_1^\infty \frac{\phi^2(u)}{u} \exp(-\phi^2(u))\, du$ converges or diverges.

• **Central limit theorem and functional central limit theorem** ∎

Remark 1.10 (a) Berger [4] gives a unified approach to the ASIP for weakly dependent sequences of random variables with values in a real separable Banach space, but with error

term $o(\sqrt{N\log\log N})$. It follows from Berger [4, Corollary 4.1, part A.5] and Melbourne & Nicol [30] that the Banach space-valued ASIP formulated in [4, Theorem 3.2] holds for all dynamical systems considered in this paper. In particular, the $\mathbb{R}^d$-valued ASIP holds with error term $o(\sqrt{N\log\log N})$. This error term suffices for the functional LIL, but is inadequate for the upper and lower class refinements and for the (functional) central limit theorem; whereas the error term established in this paper suffices. Indeed this was the original motivation of Jain *et al.* [26] to improve the error term in Strassen's scalar ASIP.

(b) The $\mathbb{R}^d$-valued functional central limit theorem, being a distributional result, can be proved directly under the more general condition $R \in L^2$ in Theorem 1.6: reduce as in this paper to the setting in Section 3 and then apply the method of [21, Section 3.3].

We end this section by discussing briefly the probabilistic methods used in this paper. Strassen's original proof of the scalar ASIP for IIDs and martingales [40, 41] relies heavily on the Skorokhod embedding theorem for scalar stochastic processes. This method was extended to *weakly dependent* sequences of random variables by a number of authors, using *blocking arguments* to reduce to the martingale case, see [38]. In particular, Philipp & Stout [38, Theorem 7.1] formulated a version of the scalar ASIP which is particularly useful for dynamical systems [24, 17, 30].

Attempts to extend Strassen's proof to the $\mathbb{R}^d$-valued situation were only partially successful [27], and the same is true for the completely different *quantile transform method* of Csörgö & Révész [15]. Eventually, Berkes & Philipp [5] introduced a third method which works in any number of dimensions, and the applicability of this method was extended to weakly dependent sequences by Kuelbs & Philipp [28].

The remainder of the paper is organised as follows. In Section 2, we combine the blocking argument in [38] with the results of [5, 28] to prove a vector-valued ASIP for $\mathbb{R}^d$-valued random variables satisfying certain hypotheses. In Section 3, we verify these hypotheses for Gibbs-Markov maps and derive Theorem 1.3 as a consequence. In Section 4, we first prove the ASIP for nonuniformly expanding maps and then prove Theorems 1.6 and 1.7. We also list numerous other situations to which our results apply, and we mention some open problems regarding time-one maps of flows.

2 A vector-valued ASIP for functions of mixing sequences

In this section, we prove a vector-valued ASIP for $\mathbb{R}^d$-valued random variables satisfying certain hypotheses. This is the vector-valued analogue of [38, Theorem 7.1] though with hypotheses tailored to the dynamical systems setting. (A result of this type is hinted at in Kuelbs & Philipp [28], but it is necessary to work through the details to determine the hypotheses, which were left unstated. In any case, the estimates in (2.3) and (2.5) are not

so natural in the probabilistic setting in [28], and partly account for our strong error term.)

2.1 Statement of the ASIP

Let $\xi_1, \xi_2, \ldots$ be a sequence of real-valued random variables and let $\mathcal{F}_a^b = \sigma\{\xi_n; a \le n \le b\}$. We assume the strong-mixing condition

$$|P(AB) - P(A)P(B)| \le C\tau^n \quad \text{for all } A \in \mathcal{F}_1^k \text{ and } B \in \mathcal{F}_{k+n}^\infty. \tag{2.1}$$

Let $p > 2$, and let η_n be a strictly stationary sequence of $\mathcal{F}_n^\infty$-measurable $\mathbb{R}^d$-valued random variables satisfying

$$E\eta_n = 0 \quad \text{and} \quad |\eta_n|_p = C, \tag{2.2}$$

and the (backwards) Burkholder-type inequality

$$\Big|\max_{1\le\ell\le N} \big|\textstyle\sum_{n=\ell}^N \eta_n\big|\Big|_p \le CN^{\frac{1}{2}}. \tag{2.3}$$

Define $\eta_{\ell n} = E(\eta_n|\mathcal{F}_n^{n+\ell})$. We require that

$$|\eta_n - \eta_{\ell n}|_p \le C\tau^\ell. \tag{2.4}$$

Let Σ be a symmetric positive semidefinite $d \times d$ covariance matrix. Given $u \in \mathbb{R}^d$, define $f_N(u) = E\exp(i\langle u, \sum_{n\le N}\eta_n/\sqrt{N}\rangle)$ and $g(u) = \exp(-\frac{1}{2}\langle u, \Sigma u\rangle)$. Assume that there exists $\epsilon > 0$ such that

$$|f_N(u) - g(u)| \le CN^{-\frac{1}{2}} \text{ for all } |u| \le \epsilon N^{\frac{1}{2}}. \tag{2.5}$$

Theorem 2.1 *Assume conditions* (2.1)–(2.5). *Let* $\beta = \frac{\frac{1}{p}+2d+3}{4d+7} \in \left[\frac{2d+3}{4d+7}, \frac{1}{2}\right)$, *and let* $\epsilon > 0$. *Then there is a d-dimensional Brownian motion $W(t)$ with covariance matrix Σ such that* $\sum_{n\le N}\eta_n = W(N) + O(N^{\beta+\epsilon})$ *a.e.*

Remark 2.2 (a) For $d = 1$, we obtain under similar hypotheses, but using a different method, the improved error estimate $\beta = \frac{1}{2p} + \frac{1}{4}$ for $2 < p \le 4$ and $\beta = \frac{3}{8}$ for $p \ge 4$. See Appendix A.

(b) It is evident from the proof of Theorem 2.1 that the exponential rates in (2.1) and (2.4) can be replaced by sufficiently high polynomial rates. Further relaxing of the assumptions is possible at the cost of obtaining a weaker estimate in Theorem 2.1.

2.2 Preliminaries

The following result of [16, 43] is stated as [38, Lemma 7.2.1]

Lemma 2.3 *Let $\mathcal{F}, \mathcal{G}$ be σ-fields and $\beta \geq 0$ such that $|P(AB) - P(A)P(B)| \leq \beta$ for all $A \in \mathcal{F}$, $B \in \mathcal{G}$. Let $p, q, r > 1$ satisfy $\frac{1}{p} + \frac{1}{r} + \frac{1}{s} = 1$. Suppose that $\xi \in L^r(\mathcal{F})$, $\eta \in L^s(\mathcal{G})$. Then $|E(\xi\eta) - E(\xi)E(\eta)| \leq 10\beta^{\frac{1}{p}}\|\xi\|_r\|\eta\|_s$.*

The next result is due to Dvoretsky [19], see [28, Lemma 2.2].

Lemma 2.4 *Let $\mathcal{F}, \mathcal{G}$ be σ-fields. Let ξ be a complex-valued $\mathcal{F}$-measurable random variable with $|\xi| \leq 1$. Then $E|E(\xi|\mathcal{G}) - E\xi| \leq 2\pi \sup_{A\in\mathcal{F},B\in\mathcal{G}} |P(AB) - P(A)P(B)|$.*

The following Gal-Koksma strong law [22] is stated in [38, Theorem A1].

Lemma 2.5 *Let ξ_j be a sequence of random variables with $E\xi_j = 0$, and let $q > 0$. Suppose that $E|\sum_{j=m}^{n} \xi_j|^2 \leq n^q - m^q$ for all $n \geq m \geq 1$. For any $\epsilon > 0$, $\sum_{j=1}^{M} \xi_j \ll M^{\frac{q}{2}+\epsilon}$ a.e.*

2.3 Introduction of the blocks

Fix $Q > \alpha > 0$. Define random variables $y_1, z_1, y_2, z_2 \ldots$ consisting of sums of consecutive $\eta_{\ell(n),n}$ where the j'th blocks y_j and z_j consist of $[j^Q]$ and $[j^\alpha]$ such terms respectively, and throughout the j'th blocks $\ell(n) = [\frac{1}{2}j^\alpha]$.

In other words, $y_j = \sum_n \eta_{\ell,n}$, where $\ell = [\frac{1}{2}j^\alpha]$ and the sum ranges over $\sum_{i=1}^{j-1}([i^Q]+[i^\alpha]) < n \leq \sum_{i=1}^{j-1}([i^Q] + [i^\alpha]) + [j^Q]$. Similarly for z_j.

Let $\mathcal{L}_a^b = \sigma\{y_j; a \leq j \leq b\}$ and $\tilde{\mathcal{L}}_a^b = \sigma\{z_j; a \leq j \leq b\}$.

Lemma 2.6 *There exists (a modified) $\tau \in (0, 1)$ such that for all $k, n \geq 1$,*

$$|P(AB) - P(A)P(B)| \ll \tau^{(k+n)^\alpha} \quad \text{for all } A \in \mathcal{L}_1^k \text{ and } B \in \mathcal{L}_{k+n}^\infty.$$

The same is true for all $A \in \tilde{\mathcal{L}}_1^k$ and $B \in \tilde{\mathcal{L}}_{k+n}^\infty$.

Proof Note that $\mathcal{L}_1^k$ is defined using $y_1, \ldots, y_k$ which are defined using $\eta_{\ell n}$ with $\ell \leq [\frac{1}{2}k^\alpha]$, $n \leq \sum_{i=1}^{k-1}([i^Q] + [i^\alpha]) + [k^Q]$. This involves conditioning on ξ_n with $n \leq \sum_{i=1}^{k-1}([i^Q] + [i^\alpha]) + [k^Q] + [\frac{1}{2}k^\alpha]$. Similarly for $\mathcal{L}_{k+n}^\infty$ and we obtain

$$\mathcal{L}_1^k \subset \mathcal{F}_1^{\sum_{i=1}^{k-1}([i^Q]+[i^\alpha])+[k^Q]+[\frac{1}{2}k^\alpha]}, \quad \mathcal{L}_{k+n}^\infty \subset \mathcal{F}^\infty_{\sum_{i=1}^{k+n-1}([i^Q]+[i^\alpha])+1}.$$

Hence $|P(AB) - P(A)P(B)| \leq \tau^N$ where $N = \sum_{i=k+1}^{k+n-1}([i^Q] + [i^\alpha]) + [k^\alpha] - [\frac{1}{2}k^\alpha] + 1$. For all $k, n \geq 1$, we compute that $N \gg (k + n)^\alpha$ as required for the first statement. (Note that the

details for the cases $n = 1$ and $n \geq 2$ are slightly different.) The second statement is proved in the same way. ∎

For $N \geq 1$, let y_{M_N}, z_{M_N} be the pair of blocks that contains $\eta_{\ell(N),N}$. Write

$$y_{M_N} + z_{M_N} = \textstyle\sum_{j=P_{M_N-1}+1}^{P_{M_N}} \eta_{\ell j}, \quad \ell = [\tfrac{1}{2}M_N^\alpha].$$

In particular, $P_{M_N-1} < N \leq P_{M_N}$, and $P_{M_N} - P_{M_N-1} = [M_N^Q] + [M_N^\alpha] \sim M_N^Q$. It is immediate that

Proposition 2.7 *Writing $M = M_N$, we have $M^{1+Q} \sim \sum_{j \leq M} j^Q \sim N$. In particular, $P_M - P_{M-1} \sim N^{Q/(1+Q)}$.* ∎

Proposition 2.8 $\sum_{n \geq 1} |\eta_n - \eta_{\ell n}|_p < \infty$.

Proof Focusing on the M'th block, and applying (2.4), we obtain $\sum_{P_{M-1} < n \leq P_M} |\eta_n - \eta_{\ell(n),n}|_p \ll M^Q \tau^{\frac{1}{2}M^\alpha}$ which is summable. ∎

Proposition 2.9 $|y_j|_p \ll j^{\frac{1}{2}Q}$ *and* $|z_j|_p \ll j^{\frac{1}{2}\alpha}$.

Proof Write $y_j = \sum^* \eta_{\ell n}$ where $\sum^* = \sum_{n=a_j+1}^{a_j+[j^Q]}$. By Proposition 2.8, (2.3) and stationarity, $|y_j|_p \leq |\sum^*(\eta_{\ell n} - \eta_n)|_p + |\sum^* \eta_n|_p \ll 1 + j^{\frac{1}{2}Q} \ll j^{\frac{1}{2}Q}$. Similarly for z_j. ∎

2.4 Approximation result

Theorem 2.10 *Let $\beta = \max\{\frac{1}{p} + \frac{1}{2}Q, \frac{1}{2}\}/(1+Q)$. For any $\epsilon > 0$, there exists $\alpha > 0$ such that $\sum_{n \leq N} \eta_n - \sum_{j \leq M_N} y_j \ll N^{\beta+\epsilon}$ a.e.*

Begin by writing

$$\sum_{n \leq N} \eta_n - \sum_{j \leq M_N} y_j = \Big(\sum_{n \leq P_{M_N}} \eta_n - \sum_{j \leq M_N} (y_j + z_j) \Big) - \sum_{n=N+1}^{P_{M_N}} \eta_n + \sum_{j \leq M_N} z_j.$$

In the next three lemmas, we estimate these three terms (following [38, Lemmas 7.3.2, 7.3.3, 7.3.4]). The result follows by combining these estimates.

Lemma 2.11 $\sum_{n \leq P_{M_N}} \eta_n - \sum_{j \leq M_N}(y_j + z_j) \ll 1$ *a.e.*

Proof By Proposition 2.8, $\sum_{n \leq \infty} |\eta_n - \eta_{\ell n}| < \infty$ a.e. Hence $|\sum_{n \leq P_M} \eta_n - \sum_{j \leq M}(y_j + z_j)| = |\sum_{n \leq P_M}(\eta_n - \eta_{\ell n})| \leq \sum_{n \leq \infty} |\eta_n - \eta_{\ell n}| \ll 1$ a.e. ∎

Lemma 2.12 *Let $\beta = (\frac{1}{2}+\frac{1}{2}\alpha)/(1+Q)$. For any $\epsilon > 0$, $\sum_{j\le M_N} z_j \ll N^{\beta+\epsilon}$ a.e.*

Proof We have $|z_j|_p \ll j^{\frac{1}{2}\alpha}$ and so $\sum_m^n Ez_j^2 \ll \sum_m^n j^\alpha \ll n^{1+\alpha} - m^{1+\alpha}$. By Lemmas 2.3 and 2.6 (with $\tilde\tau = \tau^\epsilon$ where $\epsilon = 1-2/p$), for all $i<j$,

$$|Ez_iz_j| \ll |z_i|_p|z_j|_p\tilde\tau^{j^\alpha} \le (i^\alpha\tilde\tau^{i^\alpha} j^\alpha\tilde\tau^{j^\alpha})^{\frac{1}{2}}$$

which is summable over $(i,j)\in\mathbb{N}^2$. We have shown that $E(\sum_{j=m}^n z_j)^2 \ll n^{1+\alpha}-m^{1+\alpha}$, for all $1\le m\le n$. By Lemma 2.5, $\sum_{j\le M} z_j \ll M^\gamma$ a.e. for $\gamma > \frac{1}{2}(1+\alpha)$, and the result follows from Proposition 2.7. ∎

Lemma 2.13 *Let $\beta = (\frac{1}{p}+\frac{1}{2}Q)/(1+Q)$. For any $\epsilon>0$, $\sum_{n=N+1}^{P_{M_N}} \eta_n \ll N^{\beta+\epsilon}$ a.e.*

Proof Let $A_M = \max_{P_{M-1}+1\le N\le P_M} |\sum_{n=N+1}^{P_M}\eta_n|$. By (2.3) and stationarity, $|A_M|_p \ll (P_M - P_{M-1})^{\frac{1}{2}} \ll M^{\frac{1}{2}Q}$. Hence

$$P(A_M > M^\gamma) = P(A_M^p > M^{p\gamma}) \ll M^{-p(\gamma-\frac{1}{2}Q)},$$

which is summable provided $\gamma > \frac{1}{p}+\frac{1}{2}Q$. By Borel-Cantelli, $A_M \ll M^\gamma$ a.e. and the result follows from Proposition 2.7. ∎

2.5 Proof of Theorem 2.1

We follow the argument of Kuelbs & Philipp [28] which extends Berkes & Philipp [5]. Let $X_j = [j^Q]^{-\frac{1}{2}}y_j$. Note that $\mathcal{L}_1^j$ is an increasing sequence of σ-fields such that X_j is $\mathcal{L}_1^j$-measurable.

Proposition 2.14 *Let $\gamma\in(0,\frac{1}{2}Q)$. There exists $\epsilon>0$ such that $E\big|E(\exp(i\langle u,X_j\rangle)|\mathcal{L}_1^{j-1}) - \exp(-\frac{1}{2}\langle u,\Sigma u\rangle)\big| \le C'j^{\gamma-\frac{1}{2}Q}$ for all $u\in\mathbb{R}^d$ satisfying $|u|\le\epsilon j^\gamma$.*

Proof Let $f_N(u) = E\exp(i\langle u, \sum_{n\le N}\eta_n/\sqrt{N}\rangle)$, $g(u) = \exp(-\frac{1}{2}\langle u,\Sigma u\rangle)$, and write

$$\begin{aligned}&E\{\exp(i\langle u,X_j\rangle)|\mathcal{L}_1^{j-1}\} - g(u) = \big(E\{\exp(i\langle u,X_j\rangle)|\mathcal{L}_1^{j-1}\} - E\exp(i\langle u,X_j\rangle)\big)\\ &+\Big(E\exp(i\langle u,[j^Q]^{-\frac{1}{2}}y_j\rangle) - E\exp(i\langle u,[j^Q]^{-\frac{1}{2}}\textstyle\sum_{n\le[j^Q]}\eta_n\rangle)\Big) + \big(f_{[j^Q]}(u)-g(u)\big)\\ &= I + II + III.\end{aligned}$$

Using Lemmas 2.4 and 2.6, $E|I| \ll \tau^{j^\alpha}$. Also, III is estimated by (2.5) so it remains to estimate II. Write $y_j = \sum^* \eta_{\ell n}$ where $\sum^* = \sum_{n=a_j+1}^{a_j+[j^Q]}$. By stationarity and Proposition 2.8,

$$\begin{aligned}
|II| &= |E\Big(\exp(i\langle u, [j^Q]^{-\frac12}\textstyle\sum^*\eta_{\ell n}\rangle) - \exp(i\langle u, [j^Q]^{-\frac12}\sum^*\eta_n\rangle)\Big)| \\
&\le |\exp(i\langle u, [j^Q]^{-\frac12}\textstyle\sum^*(\eta_{\ell n} - \eta_n)\rangle) - 1|_1 \le |\langle u, [j^Q]^{-\frac12}\sum^*(\eta_{\ell n} - \eta_n)\rangle|_1 \\
&\le \epsilon j^\gamma [j^Q]^{-\frac12} |\textstyle\sum_{n\ge1}(\eta_{\ell n} - \eta_n)|_1 \ll j^{\gamma - \frac12 Q}
\end{aligned}$$

as required. ∎

Proposition 2.15 *Let G be the distribution function of $N(0,\Sigma)$. Then $G\{u : |u| > T\} \le e^{-DT^2}$.*

Proof This is a straightforward calculation, see for example [5, p. 43]. ∎

Let $\lambda_j = C' j^{\gamma - \frac12 Q}$, $T_j = \epsilon j^\gamma$, where $\gamma \in (0, \frac12 Q)$ is chosen below. By Propositions 2.14 and 2.15, we have

$$\begin{aligned}
&E|E\{\exp(i\langle u, X_j\rangle)|\mathcal{L}_1^{j-1}\} - g(u)| \le \lambda_j \text{ for all } |u| \le T_j, \\
&G\{u : |u| > \tfrac14 T_j\} \le \delta_j,
\end{aligned}$$

where $\delta_j = e^{-D' j^{2\gamma}}$. These are the hypotheses of [5, Theorem 1]. Defining

$$\alpha_j = 16d\, T_j^{-1} \log T_j + 4\lambda_j^{\frac12} T_j^d + \delta_j,$$

as in [5], we have $\alpha_j \ll j^{-\gamma} \log j + j^{(d+\frac12)\gamma - \frac14 Q}$, which is summable provided $1 < \gamma < \frac{\frac14 Q - 1}{d + \frac12}$. We take γ slightly larger than 1 and Q slightly larger than $4d+6$ so that $\alpha_j \ll j^{-(1+\epsilon)}$.

Applying [5, Theorem 1], we conclude that (passing to a richer probability space) there is a sequence of i.i.d. random variables Y_j with distribution $N(0,\Sigma)$ such that

$$|X_j - Y_j| \ll j^{-(1+\epsilon)} \quad \text{a.e.}$$

Let $W(t)$ be a Brownian motion with covariance Σ and define $Y_j^* = [j^Q]^{-\frac12}(W(h_j) - W(h_{j-1}))$ where $h_j = \sum_{n=1}^j [n^Q] \sim j^{1+Q}$. Then $\{Y_j\} =_d \{Y_j^*\}$ and without loss (after passing to a richer probability space), $Y_j = Y_j^*$. We have

$$\begin{aligned}
\textstyle\sum_{j\le M} y_j &= \textstyle\sum_{j\le M} [j^Q]^{\frac12} X_j = \sum_{j\le M} [j^Q]^{\frac12}(X_j - Y_j) + \sum_{j\le M} W(h_j) - W(h_{j-1}) \\
&= \textstyle\sum_{j\le M} [j^Q]^{\frac12}(X_j - Y_j) + W(h_M).
\end{aligned}$$

Now

$$\sum_{j\le M}[j^Q]^{\frac12}(X_j-Y_j)\ll\sum_{j\le M}j^{\frac12 Q}\alpha_j\ll\sum_{j\le M}j^{\frac12 Q-1}\ll M^{\frac12 Q}\ll N^{\frac12 Q/(1+Q)}.$$

If $h_M>N$, then $h_M-N\le P_M-P_{M-1}\ll M^Q$, whereas if $h_M<N$ then $N-h_M<P_M-h_M=\sum_{j\le M}[j^\alpha]\ll M^{1+\alpha}$. Hence $h_M-N\ll N^{\max\{Q,1+\alpha\}/(1+Q)}$. Taking α small, we obtain $W(h_M)=W(N)+O(N^{\max\{\frac12 Q,\frac12\}/(1+Q)+\epsilon})$. Combining these estimates with Theorem 2.10 we obtain

$$\sum_{n\le N}\eta_n=\sum_{n\le N}\eta_n-\sum_{j\le M}y_j+\sum_{j\le M}y_j=W(N)+O(N^{\max\{\frac1p+\frac12 Q,\frac12\}/(1+Q)+\epsilon}).$$

Taking Q slightly larger than $4d+6$ yields the required result. ∎

3 ASIP for Gibbs-Markov maps

In this section we prove the ASIP for weighted Lipschitz $\mathbb{R}^d$-valued observables of Gibbs-Markov maps. Roughly speaking, these are uniformly expanding maps with countably many inverse branches and good distortion properties, and have been studied extensively in [1]. We derive Theorem 1.3 as a consequence.

3.1 Gibbs-Markov maps

Let (Λ,m) be a Lebesgue space with a countable measurable partition α. Without loss, we suppose that all partition elements $a\in\alpha$ have $m(a)>0$. Recall that a measure-preserving transformation $F:\Lambda\to\Lambda$ is a *Markov map* if Fa is a union of elements of α and $F|_a$ is injective for all $a\in\alpha$. Define α' to be the coarsest partition of Λ such that Fa is a union of atoms in α' for all $a\in\alpha$. (So α' is a coarser partition than α.) If $a_0,\dots,a_{n-1}\in\alpha$, we define the n-cylinder $[a_0,\dots,a_{n-1}]=\cap_{i=0}^{n-1}F^{-i}a_i$. It is assumed that F and α separate points in Λ (if $x,y\in\Lambda$ and $x\ne y$, then for n large enough there exist distinct n-cylinders that contain x and y).

Let $0<\beta<1$. We define a metric d_β on Λ by $d_\beta(x,y)=\beta^{s(x,y)}$ where $s(x,y)$ is the greatest integer $n\ge0$ such that x,y lie in the same n-cylinder. Define $g=JF^{-1}=\frac{dm}{d(m\circ F)}$ and set $g_k=g\,g\circ F\cdots g\circ F^{k-1}$.

A Markov map F is *topologically mixing* if for all $a,b\in\alpha$ there exists $N\ge1$ such that $F^na\cap b\ne\emptyset$ for all $n\ge N$. A Markov map F is *Gibbs-Markov* if

(i) *Big images property:* There exists $c>0$ such that $m(Fa)\ge c$ for all $a\in\alpha$.

(ii) *Distortion:* $\log g|_a$ is Lipschitz with respect to d_β for all $a\in\alpha'$.

Let α_0^{k-1} denote the partition of Λ into length k cylinders $a=[a_0,\dots,a_{k-1}]$. The following result of [2] is stated explicitly in [30, Lemma 2.4(b)].

Lemma 3.1 *Let F be a topologically mixing Gibbs-Markov map. Then $\big|m(a\cap F^{-(N+k)}b) - m(a)m(b)\big| \le C\tau^N m(a)m(b)^{1/2}$ for all $a\in\alpha_0^{k-1}$ and all measurable b.* ∎

3.2 Weighted Lipschitz observations

Let $p\in[1,\infty)$. We fix a sequence of weights $R(a)>0$ satisfying $|R|_p = (\sum_{a\in\alpha} m(a)R(a)^p)^{1/p} < \infty$. Given $\Phi:\Lambda\to\mathbb{R}$ continuous, define $|\Phi|_\beta$ to be the Lipschitz constant of Φ with respect to the metric d_β. Let $\|\Phi\|_\infty = \sup_{a\in\alpha}|\Phi 1_a|_\infty/R(a)$, $\|\Phi\|_\beta = \sup_{a\in\alpha}|\Phi 1_a|_\beta/R(a)$. Let $\mathcal{B}$ consist of the space of weighted Lipschitz functions with $\|\Phi\| = \|\Phi\|_\infty + \|\Phi\|_\beta < \infty$. Note in particular that $R\in\mathcal{B}$ and $\|R\|=1$. We have the embeddings $\mathrm{Lip}\subset\mathcal{B}\subset L^p\subset L^1$, where Lip is the space of (globally) Lipschitz functions.

Lemma 3.2 *Let $\Phi\in\mathcal{B}$ with $\int_\Lambda\Phi = 0$. Then $|\Phi - E(\Phi|\alpha_0^{k-1})|_p \le \|\Phi\|_\beta |R|_p\beta^k$ for all $k\ge1$.*

Proof (cf. [30, Lemma 2.4(a)]) Note that $E(\Phi|\alpha_0^{k-1})$ is constant on partition elements $a\in\alpha_0^{k-1}$ with value $\frac{1}{m(a)}\int_a\Phi\,dm$, and that $|\Phi 1_a - \frac{1}{m(a)}\int_a\Phi\,dm|_\infty \le |\Phi 1_a|_\beta\,\mathrm{diam}_\beta(a) \le \|\Phi\|_\beta R(a)\beta^k$. Hence, $|\Phi - E(\Phi|\alpha_0^{k-1})|_p^p \le (\|\Phi\|_\beta\beta^k)^p\sum_{a\in\alpha_0^{k-1}} R(a)^p m(a) = (\|\Phi\|_\beta\beta^k|R|_p)^p$. ∎

3.3 A maximal inequality

Given a measure-preserving transformation $F:\Lambda\to\Lambda$ of a probability space (Λ,m), the transfer (Perron-Frobenius) operator $L:L^1\to L^1$ is defined by $\int_\Lambda L\Phi\,\Psi\,dm = \int_\Lambda \Phi\,\Psi\circ F\,dm$ for all $\Phi\in L^1$, $\Psi\in L^\infty$. This restricts to an operator on L^p, $1\le p\le\infty$.

Lemma 3.3 *Let $\Phi\in L^p(\Lambda)$, $1\le p<\infty$ with $L\Phi=0$. Then $\big|\max_{0\le\ell\le N-1}\big|\sum_{n=\ell}^N \Phi\circ F^n\big|\,\big|_p \le CN^{\frac12}$.*

Proof Note that $L = E(\cdot|F^{-1}\mathcal{M})$ where $\mathcal{M}$ is the underlying σ-algebra. By hypothesis the sequence $\{\Phi\circ F^n; n\ge0\}$ is a reverse martingale difference sequence. Passing to the natural extension we obtain an L^p martingale difference sequence $\{w_n; n\in\mathbb{Z}\}$ such that $\Phi\circ F^n = w_{-n}$. By Burkholder's inequality [8][1], we have $|\max_{1\le k\le N}|\sum_{n=0}^k w_n|\,|_p \le CN^{\frac12}$. Setting $\ell = N-k$ and using stationarity,

$$\max_{0\le\ell\le N-1}|\textstyle\sum_\ell^N \Phi\circ F^n| =_d \max_{0\le\ell\le N-1}|\sum_{-N+\ell}^0 \Phi\circ F^n| = \max_{1\le k\le N}|\sum_0^k w_n|,$$

proving the result. ∎

[1] This follows from [8, eqns (1.4) and (3.3)] and is stated explicitly in [37, Eq. 1].

3.4 Quasicompactness and the central limit theorem

Let $F : \Lambda \to \Lambda$ be a topologically mixing Gibbs-Markov map with transfer operator $L : L^1 \to L^1$. It is well-known [1, 30] that L restricts to a bounded operator on weighted Lipschitz observables $\Phi \in \mathcal{B}$ and $L(\mathcal{B}) \subset \mathrm{Lip}$. Moreover $L : \mathcal{B} \to \mathcal{B}$ is *quasicompact*: $L1 = 1$ and the spectral radius of L restricted to $\mathcal{B}_0 = \{\Phi \in \mathcal{B} : \int_\Lambda \Phi \, dm = 0\}$ is strictly less than 1.

We define $\mathcal{B}^d$ to the be the space of $\mathbb{R}^d$ weighted Lipschitz observables, so $\Phi = (\Phi_1, \dots, \Phi_d) \in \mathcal{B}^d$ if and only if $\Phi_i \in \mathcal{B}$ for $i = 1, \dots, d$. Similarly, we define $\mathcal{B}_0^d$. We suppress the superscript for spaces such as L^p and Lip relying on the context.

Proposition 3.4 *Suppose that $\Phi \in \mathcal{B}_0^d$. Then there exists $\Psi \in \mathcal{B}_0^d$ and $\chi \in L^\infty$ such that $\Phi = \Psi + \chi \circ F - \chi$ and $L\Psi = 0$.*

Proof (cf. [30, Proof of Corollary 2.3(c)]) Define $\chi = \sum_{j=1}^\infty L^j \Phi$. This converges in $\mathcal{B}_0^d$ since the spectral radius of L is less than 1. Since $L(\mathcal{B}^d) \subset \mathrm{Lip}$, we have $\chi \in L^\infty$. By construction, $L\Psi = 0$. ∎

Suppose that $p \geq 2$. Let $\Phi \in \mathcal{B}_0^d \subset L^2$ and assume that $L\Phi = 0$. Let $S_N = \sum_{n \leq N} \Phi \circ F^n$ and form the $d \times d$ matrix $S_N S_N^T$. We define the covariance matrix $\Sigma = \frac{1}{N} \int_\Lambda S_N S_N^T \, dm = \int_\Lambda \Phi \Phi^T \, dm$.

Lemma 3.5 *There exists $\epsilon > 0$ such that*

$$\textstyle\int_\Lambda \exp(i\langle u, S_N\rangle N^{-\frac{1}{2}}) \, dm - \exp(-\frac{1}{2}\langle u, \Sigma u\rangle) = O(N^{-\frac{1}{2}})$$

uniformly for $u \in \mathbb{R}^d$ satisfying $|u| \leq \epsilon N^{\frac{1}{2}}$.

Proof We follow a standard argument establishing the central limit theorem with error term for systems with quasicompact transfer operator (see [36, Theorem 4.13] and references therein). Let S^{d-1} denote the unit sphere in $\mathbb{R}^d$. Given $u \in \mathbb{R}^d$, write $u = tv$ where $t \geq 0$ and $v \in S^{d-1}$. Define the twisted transfer operator $L_u : \mathcal{B} \to \mathcal{B}$ by $L_u \Psi = L(e^{i\langle u, \Phi\rangle}\Psi)$. Recall that 1 is an isolated eigenvalue for L. For u small, the spectral radius of L_u is $\exp P(u)$ where P is analytic, $P(0) = 0$. Moreover [36, p. 66]

$$P(u) = -\tfrac{1}{2}\langle u, \Sigma u\rangle - iP_3(v)t^3 + P_4(v, t)t^4,$$

where $P_3(v) \in \mathbb{R}$, $P_4(v, t) \in \mathbb{C}$ are analytic, and there exists $\epsilon > 0$ such that [36, p. 67]

$$\exp(NP(uN^{-\frac{1}{2}})) - \exp(-\tfrac{1}{2}\langle u, \Sigma u\rangle)(1 - iP_3(v)t^3N^{-\frac{1}{2}}) = O(N^{-1})$$

uniformly for $|u| \leq \epsilon N^{\frac{1}{2}}$.

Now $\int_\Lambda \exp(i\langle u, S_N\rangle N^{-\frac{1}{2}})\, dm = \int_\Lambda (L_{uN^{-\frac{1}{2}}})^N 1\, dm$, and since the leading eigenvalue of L_u is isolated there exists $\gamma \in (0,1)$ such that

$$\int_\Lambda \exp(i\langle u, S_N\rangle N^{-\frac{1}{2}})\, dm - \exp(NP(uN^{-\frac{1}{2}})) \ll \gamma^N,$$

uniformly for $|u| \le \epsilon N^{\frac{1}{2}}$. This completes the proof. ∎

3.5 Statement and proof of ASIP for Gibbs-Markov maps

Theorem 3.6 *Suppose that $F : \Lambda \to \Lambda$ is a topologically mixing Gibbs-Markov map. Define the Banach space $\mathcal{B}^d$ corresponding to weights $R \in L^p$ where $p > 2$. Suppose that $\Phi : \Lambda \to \mathbb{R}^d$ is a mean zero observable in $\mathcal{B}^d$ with partial sums $S_N = \sum_{n=1}^N \Phi \circ F^j$. Then the conclusion of Theorem 1.6 is valid.*

Proof By Proposition 3.4, $S_N = \sum_{n=1}^N \Psi \circ F^j + O(1)$ a.e. where $L\Psi = 0$. Hence without loss we may suppose from the outset that $L\Phi = 0$.

Define $\eta_n = \Phi \circ F^n$ and $\xi_n = a_n$. Then $\eta_n = \Phi(\xi_n, \xi_{n+1}, \ldots)$. We verify the hypotheses of Theorem 2.1.

Hypothesis (2.1) follows from Lemma 3.1 and (2.2) is immediate. The remaining hypotheses follow from Lemmas 3.3, 3.2 and 3.5 respectively.

This completes the proof for $d \ge 2$. The improved estimate for $d = 1$ follows from Theorem A.9. (One hypothesis is different, but it was verified in [30].) ∎

Proof of Theorem 1.3 This reduces by standard techniques to a two-sided and then one-sided subshift of finite type. The latter is a special case of a Gibbs-Markov map with finite alphabet, hence $R \in L^\infty$. Theorem 1.3 follows from Theorem 3.6 with $p = \infty$. ∎

4 Applications to nonuniformly hyperbolic systems

In this section, we prove the vector-valued ASIP for large classes of nonuniformly hyperbolic systems. In Subsection 4.1, we consider nonuniformly expanding systems. In Subsection 4.2, we consider nonuniformly hyperbolic systems, proving Theorems 1.6 and 1.7. Some open problems are described in Subsection 4.3.

4.1 Nonuniformly expanding systems

Let (M, d) be a locally compact separable bounded metric space with Borel probability measure η and let $f : M \to M$ be a nonsingular transformation for which η is ergodic. Let $\Lambda \subset M$ be a measurable subset with $\eta(\Lambda) > 0$. We suppose that there is an at most

countable measurable partition $\{\Lambda_j\}$ with $\eta(\Lambda_j) > 0$, and that there exist integers $R_j \geq 1$, and constants $\lambda > 1$; $C > 0$ and $\gamma \in (0,1)$ such that for all j,

(1) $f^{R_j} : \Lambda_j \to \Lambda$ is a (measure-theoretic) bijection.

(2) $d(f^{R_j}x, f^{R_j}y) \geq \lambda d(x,y)$ for all $x, y \in \Lambda_j$.

(3) $d(f^k x, f^k y) \leq C d(f^{R_j}x, f^{R_j}y)$ for all $x, y \in \Lambda_j$, $k < R_j$.

(4) $g_j = \frac{d(\eta|_{\Lambda_j} \circ (f^{R_j})^{-1})}{d\eta|_\Lambda}$ satisfies $|\log g_j(x) - \log g_j(y)| \leq C d(x,y)^\gamma$ for almost all $x, y \in \Lambda$.

(5) $\sum_j R_j \eta(\Lambda_j) < \infty$.

A dynamical system f satisfying (1)–(5) is called *nonuniformly expanding.*

Define the *return time function* $R : \Lambda \to \mathbb{Z}^+$ by $R|_{\Lambda_j} \equiv R_j$ and the *induced map* $F : \Lambda \to \Lambda$ by $Fy = f^{R(y)}(y)$. It is well-known that there is a unique invariant probability measure m on M that is equivalent to η.

Theorem 4.1 *Let $f : M \to M$ be a nonuniformly expanding map satisfying (1)–(5) above. Assume moreover that $R \in L^p(\Lambda)$, $p > 2$. Let $\phi : M \to \mathbb{R}^d$ be a mean zero Hölder observation with partial sums $S_N = \sum_{n=1}^N \phi \circ f^j$. Then the conclusion of Theorem 1.6 is valid.*

Proof This is identical to the proof of [30, Theorem 2.9] so we just sketch the main steps. The induced map $F : \Lambda \to \Lambda$ is a topologically mixing Gibbs-Markov map with respect to the partition $\alpha = \{\Lambda_j\}$. The induced observable $\Phi : \Lambda \to \mathbb{R}^d$ given by $\Phi(y) = \sum_{\ell=0}^{R(y)-1} \phi(f^j y)$ is weighted Lipschitz and satisfies the ASIP by Theorem 3.6.

If $F : \Lambda \to \Lambda$ were the first return map, then the result would follow immediately from [32, Theorem 4.2] (see also [30, Theorem B.1]). The general result is proved by passing to a Young tower [45] which is a Markov extension of $f : M \to M$ for which F is the first return map. ∎

Remark 4.2 (a) The regularity assumption on ϕ in Theorem 4.1 can be replaced by the more general assumption that the induced observable Φ is weighted Lipschitz (with respect to the metric defined on the Gibbs-Markov system Λ).

(b) A similar result holds for nonuniformly expanding semiflows [30, Corollary 2.12].

Naturally, Theorem 4.1 includes uniformly expanding and piecewise expanding maps where the partition is finite (with $p = \infty$). Further examples of nonuniformly expanding maps to which Theorem 4.1 applies include Alves-Viana maps, Liverani-Saussol-Vaienti (Pomeau-Manneville maps), multimodal maps, and circle maps with a neutral fixed point, see [30, Section 4].

4.2 Nonuniformly hyperbolic systems

As was the case in [30], the results in this paper apply to dynamical systems that are *nonuniformly hyperbolic in the sense of Young [44]* with return time function $R \in L^p$, $p > 2$.

Let $f : M \to M$ be a diffeomorphism (possibly with singularities) defined on a Riemannian manifold (M, d). We assume from the start that f preserves a "nice" probability measure m (one of the conclusions in Young [44] is that m is a SRB measure).

Fix a subset $\Lambda \subset M$ and a family of subsets of M called "stable disks" $\{W^s\}$ that are disjoint and cover Λ. The stable disk containing x is labelled $W^s(x)$.

(A1) There is a partition $\{\Lambda_j\}$ of Λ and integers $R_j \geq 1$ such that $f^{R_j}(W^s(x)) \subset W^s(f^{R_j}x)$ for all $x \in \Lambda_j$.

Define the return time function $R : \Lambda \to \mathbb{Z}^+$ by $R|_{\Lambda_j} = R_j$ and the induced map $F : \Lambda \to \Lambda$ by $F(x) = f^{R(x)}(x)$. Form the discrete suspension map $\hat{f} : \Delta \to \Delta$ where $\hat{f}(x, \ell) = (x, \ell + 1)$ for $\ell < R(x) - 1$ and $\hat{f}(x, R(x) - 1) = (Fx, 0)$. Define a separation time $s : \Lambda \times \Lambda \to \mathbb{N}$ by defining $s(x, x')$ to be the greatest integer $n \geq 0$ such that $F^k x, F^k x'$ lie in the same partition element of Λ for $k = 0, \ldots, n$. (If x, x' do not lie in the same partition element, then we take $s(x, x') = 0$.) For general points $p = (x, \ell), p' = (x', \ell') \in \Delta$, define $s(p, q) = s(x, x')$ if $\ell = \ell'$ and $s(p, q) = 0$ otherwise. This defines a separation time $s : \Delta \times \Delta \to \mathbb{N}$. The projection $\pi : \Delta \to M$, $\pi(x, \ell) = f^\ell x$, satisfies $\pi f = \hat{f} \pi$.

(A2) There is a distinguished "unstable leaf" $W^u \subset \Lambda$ such that each stable disk intersects W^u in precisely one point, and there exist constants $C \geq 1$, $\alpha \in (0, 1)$ such that

   (i) $d(f^n x, f^n y) \leq C\alpha^n$, for all $y \in W^s(x)$, all $n \geq 0$, and

   (ii) $d(f^n x, f^n y) \leq C\alpha^{s(x,y)}$ for all $x, y \in W^u$ and all $0 \leq n < R$.

Let $\bar{\Lambda} = \Lambda / \sim$ where $x \sim x'$ if $x \in W^s(x')$ and define the partition $\{\bar{\Lambda}_j\}$ of $\bar{\Lambda}$. We obtain a well-defined return time function $R : \bar{\Lambda} \to \mathbb{Z}^+$ and induced map $\bar{F} : \bar{\Lambda} \to \bar{\Lambda}$. Let $\bar{f} : \bar{\Delta} \to \bar{\Delta}$ denote the quotient of $\hat{f} : \Delta \to \Delta$ where (x, ℓ) is identified with (x', ℓ') if $\ell = \ell'$ and $x' \in W^s(x)$. Let $\bar{\pi} : \Delta \to \bar{\Delta}$ denote the natural projection. The separation time on Δ drops down to a separation time on $\bar{\Delta}$.

(A3) The map $F : \bar{\Lambda} \to \bar{\Lambda}$ and partition $\{\bar{\Lambda}_j\}$ separate points in $\bar{\Lambda}$. (It follows that $d_\theta(p, q) = \theta^{s(p,q)}$ defines a metric on $\bar{\Delta}$ for each $\theta \in (0, 1)$.)

(A4) There exist invariant probability measures $\hat{m}$ on Δ and $\bar{m}$ on $\bar{\Delta}$ such that

   (i) $\pi : \Delta \to M$ and $\bar{\pi} : \Delta \to \bar{\Delta}$ are measure-preserving; and

   (ii) $\bar{f} : \bar{\Delta} \to \bar{\Delta}$ is nonuniformly expanding in the sense of Subsection 4.1 with induced map $\bar{F} : \bar{\Lambda} \to \Lambda$. (Conditions (2) and (3) are automatic.)

Proof of Theorem 1.6 This reduces, as in the proof of [30, Theorem 3.4], to the ASIP for the nonuniformly expanding map $\bar{f} : \bar{\Delta} \to \bar{\Delta}$ and hence follows from Theorem 4.1. ∎

Remark 4.3 Again, the regularity assumption on ϕ can be relaxed, and the result extends to nonuniformly hyperbolic flows.

Large classes of billiard maps and Lorentz flows, surveyed in [11] satisfy the vector-valued ASIP. These include dispersing billiards (with finite or infinite horizons) and the corresponding Lorentz flows (assuming finite horizons).

Proof of Theorem 1.7 By periodicity, we can consider the quotient flow on the compact manifold $M = (\mathbb{T}^2 - \Omega) \times S^1$. The Poincaré map $f : X \to X$ on the compact cross-section $X = \partial\Omega \times (\frac{-\pi}{2}, \frac{\pi}{2})$ is called the billiard map or collision map. Benedicks & Young [3] showed that f is nonuniformly hyperbolic in the sense of Young with $R \in L^p$ for all $p > 2$. By Theorem 1.6, the vector-valued ASIP holds for f with $p = \infty$. The collision time is uniformly bounded and piecewise Hölder, so it follows from [32] that the vector-valued ASIP holds for the Lorentz flow on M. Now take as an $\mathbb{R}^2$-valued observable the velocity coordinate $v : M \to S^1$. This is piecewise Hölder, and the lifted position in $\mathbb{R}^2$ is given by $q(T) = \int_0^T v \circ f_t \, dt$. Finally, nonsingularity of the covariance matrix was proved in [6]. ∎

Chernov & Zhang [12] consider three classes of billiards with slow mixing rates. The first and third classes are not covered by our results since it is shown only that $R \in L^{2-\epsilon}$. The second class of *Bunimovich-type* billiards treated in [12] satisfies $R \in L^{3-\epsilon}$. The vector-valued ASIP for such billiards (and the corresponding flows) is hence a consequence of Theorem 1.6 with error $\beta = \frac{6d+10}{12d+21}$.

As in [30], Theorem 1.6 also applies to Lozi maps and certain piecewise hyperbolic maps, Hénon-like maps and partially hyperbolic diffeomorphisms with mostly contracting direction.

A further important class of dynamical systems is *singular hyperbolic flows* [33]. Theorem 1.6 does not apply directly to such systems, but it establishes the vector-valued ASIP (with $p = \infty$) when combined with the techniques in Holland & Melbourne [25].

4.3 Open problems

Given a (non)uniformly hyperbolic flow f_t, the time-one map f_1 is only *partially hyperbolic*. For such maps the Gibbs-Markov/suspension formalism breaks down so the results in [30] and in this paper do not apply. By different methods, Melbourne & Török [31] proved that the scalar ASIP is typically valid for the time-one map of an Axiom A flow. They used rapid mixing properties to reduce to a reverse martingale difference sequence. Following [14, 21], this leads to the ASIP in reverse time and hence forwards time (since the class of such flows is closed under time reversal). Similarly, the scalar ASIP for the time-one map of the planar periodic Lorentz gas with finite horizons is typically valid (since the flow is typically rapid mixing [29] and the class of flows is closed under time reversal).

Problem 1 Prove that the vector-valued ASIP holds (at least typically) for time-one maps of Axiom A flows and/or planar periodic Lorentz gas with finite horizons.

Generally speaking, the hypotheses for a nonuniformly hyperbolic system are not time-symmetric so [14, 21, 31] fails.

Problem 2 Obtain results on the scalar ASIP for time-one maps of nonuniformly hyperbolic flows and/or singular hyperbolic flows.

Remark 4.4 (a) The Banach space-valued ASIP of [4] applies to Problem 1, with the caveats mentioned in Remark 1.10(a). In particular, the d-dimensional functional LIL is typically valid. These results do not apply to Problem 2.

(b) The (vector-valued) functional central limit theorem is typically valid in Problems 1 and 2 (combining the arguments in [21, Section 3.3] and [31]).

A Scalar ASIP with error term

In this appendix, we prove a scalar ASIP using martingale approximation and the method of Strassen [41]. This is precisely the result [38, Theorem 7.1] used in [30], but our purpose here is to obtain a better error term under assumptions appropriate for dynamical systems. This improves Theorem 2.1 when $d = 1$.

We assume the conditions of Section 2 except that (2.5) is replaced by

$$E(\textstyle\sum_{n\le N}\eta_n)^2 = N + O(N^{1/2}). \tag{A.1}$$

Define $\{y_j\}$ as in Section 2.3. In particular, Theorem 2.10 is unchanged.

Law of large numbers for y_j^2

Lemma A.1 *Let* $\gamma = \max\{\frac{\frac12 Q}{1+Q}, \frac{\frac12+\alpha}{1+Q}\}$. *Then* $\sum_{j\le M_N} Ey_j^2 = N + O(N^{\frac12+\gamma})$.

Proof (cf. [38, Lemma 7.3.5]) By (A.1), $E(\sum_{n\le N}\eta_n)^2 = a_N^2$ where $a_N^2 = N(1+O(N^{-1/2}))$. Write

$$\sum_{n\le N}\eta_n = \sum_{n\le P_M}\eta_n - \sum_{n=N+1}^{P_M}\eta_n = \sum_{n\le P_M}(\eta_n-\eta_{\ell,n}) + \sum_{j\le M} y_j + \sum_{j\le M} z_j - \sum_{n=N+1}^{P_M}\eta_n.$$

Then

$$\begin{aligned}\big\|\textstyle\sum_{j\le M} y_j\big\|_2 - a_N &= \big\|\textstyle\sum_{j\le M} y_j\big\|_2 - \big\|\textstyle\sum_{n\le N}\eta_n\big\|_2 \\ &\le \big\|\textstyle\sum_{j\le M} z_j\big\|_2 + \big\|\textstyle\sum_{n\le P_M}(\eta_n-\eta_{\ell n})\big\|_2 + \big\|\textstyle\sum_{n=N+1}^{P_M}\eta_n\big\|_2.\end{aligned}$$

By Proposition 2.8, $\|\sum_{n\le N}(\eta_n-\eta_{\ell n})\|_2 \ll 1$. By the proof of Lemma 2.12, $\|\sum_{j\le M} z_j\|_2^2 \ll M^{1+2\alpha}$ and so $\|\sum_{j\le M} z_j\|_2 \ll N^{(\frac12+\alpha)/(1+Q)}$. By stationarity and (A.1), $\|\sum_{n=N+1}^{P_M}\eta_n\|_2^2 = \|\sum_{n\le P_M-N}\eta_n\|_2^2 \ll P_M-N \ll N^{Q/(1+Q)}$. Hence $\|\sum_{j\le M} y_j\|_2 = a_N + O(N^\gamma)$ and $E(\sum_{j\le M} y_j)^2 = N + O(N^{\frac12+\gamma})$. Also, as in the proof of Lemma 2.12, $\sum_{i\ne j} Ey_iy_j \ll 1$. ∎

Corollary A.2 *Let* $\beta = \max\{\frac{\frac14+\frac12 Q}{1+Q}, \frac{\frac12+\frac12\alpha+\frac14 Q}{1+Q}\}$. *Then* $\sum_{j\le M_N} Ey_j^2 = N + O(N^{2\beta})$. ∎

Lemma A.3 *Let* $\beta = (\frac34-\frac{p}{8}+\frac12 Q)/(1+Q)$ *for* $2<p\le 4$ *(and* $\beta=(\frac14+\frac12 Q)/(1+Q)$ *for* $p>4$*). Then for any* $\epsilon>0$, $\sum_{j\le M_N} y_j^2 - Ey_j^2 \ll N^{2\beta+\epsilon}$ *a.e.*

Proof The value of ϵ below may change from line to line. Define

$$w_j = \begin{cases} y_j^2 - Ey_j^2, & |y_j^2-Ey_j^2| \le j^{1+Q+\epsilon} \\ 0, & \text{otherwise} \end{cases}$$

Note that $P(w_j \ne y_j^2-Ey_j^2) = P(|y_j^2-Ey_j^2| > j^{1+Q+\epsilon}) \le 2\|y_j\|_2^2/j^{1+Q+\epsilon} \ll j^{-(1+\epsilon)}$ which is summable, so by Borel-Cantelli w_j fails to coincide with $y_j^2-Ey_j^2$ only finitely often. Hence it suffices to estimate $\sum_{j\le M} w_j$. We do this by estimating $\sum_{j\le M}\tilde w_j$ and $\sum_{j\le M} Ew_j$ where $\tilde w_j = w_j - Ew_j$.

Again $\sum_{i\ne j} E\tilde w_i\tilde w_j \ll 1$. Also, $E\tilde w_j^2 \le |\tilde w_j^{p/2}|_1\|\tilde w_j^{2-p/2}\|_\infty \ll \|y_j\|_p^p\|w_j\|_\infty^{2-p/2} \ll j^{R-1}$, where $R = 3-\frac{p}{2}+2Q+\epsilon$. Hence $E(\sum_{j=m}^n \tilde w_j)^2 \ll n^R - m^R$. By Lemma 2.5, $\sum_{j\le M}\tilde w_j \ll M^{\frac12 R} \le N^{\frac12 R/(1+Q)}$ for any $\epsilon>0$.

Let $A = \{|y_j^2-Ey_j^2| > j^{1+Q+\epsilon}\}$. Then

$$\begin{aligned} Ew_j &= -E\Big\{(y_j^2-Ey_j^2)I_A\Big\} \ll \|y_j^2-Ey_j^2\|_{p/2}\|1_A\|_{p/(p-2)} \\ &\ll \|y_j\|_p^2\|1_A\|_{p/(p-2)} \ll j^{Q-(1+\epsilon))(p-2)/p} = j^{S-1}. \end{aligned}$$

where $S = \frac{2}{p}-\epsilon+Q$. Hence $\sum_{j\le M} Ew_j \ll M^S$. Thus, it suffices that $2\beta = \max\{\frac12 R/(1+Q), S/(1+Q)\} = \frac12 R/(1+Q)$. ∎

Martingale approximation

Lemma A.4 *Set* $\mathcal{L}_j = \mathcal{L}_1^j = \sigma\{y_1,\dots,y_j\}$. *There is a martingale difference sequence* $\{Y_j,\mathcal{L}_j\}$ *such that* $y_j = Y_j + u_j - u_{j+1}$, *where* $\|u_j\|_q \ll \tilde\tau^{j^\alpha}$ *for all* $2<q<p$.

Proof (cf. [38, Lemma 7.4.1]) Define $u_j = \sum_{k=0}^{\infty} E(y_{j+k}|\mathcal{L}_{j-1})$. We estimate $\|E(y_{j+k}|\mathcal{L}_{j-1})\|_q$ which we write for convenience as $\|E(y|\mathcal{L})\|_q$. Note that

$$\begin{aligned} E|E(y|\mathcal{L})|^q &= E\Big\{E(y|\mathcal{L})E(y|\mathcal{L})|E(y|\mathcal{L})|^{q-2}\Big\} = E\Big\{E\big\{yE(y|\mathcal{L})|E(y|\mathcal{L})|^{q-2}|\mathcal{L}\big\}\Big\} \\ &= E\big\{yE(y|\mathcal{L})E(y|\mathcal{L})^{q-2}\big\}. \end{aligned}$$

Write $\frac{1}{q}+\frac{1}{s}=1$. Then $\frac{1}{p}+\frac{1}{s}<1$, so by Lemma 2.3,

$$E|E(y|\mathcal{L})|^q \leq \|y\|_p \|E(y|\mathcal{L})^{q-1}\|_s \tilde{\tau}^{(j+k)^\alpha}.$$

Note that $\|E(y|\mathcal{L})^{q-1}\|_s = (E|E(y|\mathcal{L})|^q)^{1-\frac{1}{q}}$, and so dividing both sides by this yields $\|E(y|\mathcal{L})\|_q \leq \|y\|_p \tilde{\tau}^{(j+k)^\alpha}$. Since $\sum_{k=0}^{\infty} \tilde{\tau}^{(j+k)^\alpha} \ll j\tilde{\tau}^{j^\alpha}$, the estimate for $\|u_j\|_q$ follows (increasing $\tilde{\tau}$ slightly).

At the same time, it follows immediately that $\sum_{k=0}^{\infty} |E(y_{j+k}|\mathcal{L}_j)|_1 < \infty$ which guarantees (see eg. [38, Lemma 2.1]) that Y_j is a martingale difference sequence. ∎

Corollary A.5 *$\sum_{j\leq M_N}(y_j - Y_j) \ll 1$ a.e. and $\sum_{j\leq M_N}(y_j^2 - Y_j^2) \ll N^{\frac{1}{2}}$ a.e.*

Proof We have $\sum_{j\leq M}(y_j - Y_j) = u_1 - u_{M+1}$, and hence certainly $|\sum_{j\leq M}(y_j - Y_j)| \leq \sum_{j\geq 1}|u_j|$. By Lemma A.4, $\sum_{j\geq 1}|u_j|_1 < \infty$ so that $\sum_{j\geq 1}|u_j| < \infty$ a.e. proving the first statement.

Set $v_j = u_j - u_{j+1}$. Then $Y_j^2 - y_j^2 = v_j^2 - 2y_jv_j$. Now $E\sum_{j=1}^{\infty} v_j^2 = \sum_{j=1}^{\infty} Ev_j^2 \ll \sum_{j=1}^{\infty} \tilde{\tau}^{2j^\alpha} < \infty$ by Lemma A.4. Hence $\sum_{j\leq M} v_j^2 \leq \sum_{j=1}^{\infty} v_j^2 < \infty$ a.e. Finally, $\sum_{j\leq M} y_jv_j \leq (\sum_{j\leq M} y_j^2)^{\frac{1}{2}}(\sum_{j\leq M} v_j^2)^{\frac{1}{2}} \ll (\sum_{j\leq M} y_j^2)^{\frac{1}{2}} \ll N^{\frac{1}{2}}$ by Theorem 2.10. ∎

Lemma A.6 *Let X_k be a martingale difference sequence and $m \in (1,2]$. Suppose that $b_1 < b_2 < \cdots \to \infty$. If $\sum_{k\leq n} b_k^{-m}E|X_k|^m < \infty$, then $\sum_{k\leq n} X_k = o(b_n)$ a.e.*

Proof For $m = 2$, this is explicit in [20, p. 238]. For $m \in (1,2)$ it follows from a standard martingale result, Chow [13], combined with Kronecker's lemma; this is implicit in the proof of [38, Lemma 7.4.4]. ∎

Lemma A.7 *Let $\beta = (\frac{1}{p}+\frac{1}{2}Q)/(1+Q)$ for $2 < p \leq 4$ (and $\beta = (\frac{1}{4}+\frac{1}{2}Q)/(1+Q)$ for $p > 4$). Then for any $\epsilon > 0$, $\sum_{j\leq M_N}(E(Y_j^2|\mathcal{L}_{j-1}) - Y_j^2) \ll N^{2\beta+\epsilon}$ a.e.*

Proof Define $R_j = E(Y_j^2|\mathcal{L}_{j-1}) - Y_j^2$. Suppose that $\gamma > 2/p + Q$ and choose $q < p$ so that $\gamma > 2/q + Q$. Then

$$(j^\gamma)^{-q/2}E|R_j|^{q/2} \ll j^{-\gamma q/2}E|Y_j|^q \ll j^{-(\gamma-Q)q/2},$$

hence $\sum_{j=1}^{\infty}(j^\gamma)^{-q/2}E|R_j|^{q/2} < \infty$. Note that $\frac{q}{2} \in (1,2]$ and R_j is a martingale difference sequence, so it follows from Lemma A.6 that $\sum_{j\le M} R_j \ll M^\gamma$ and the result follows from Proposition 2.7. ∎

We now apply Strassen's martingale version of the Skorokhod embedding [41]. There exist non-negative random variables T_j such that the sequences $\big\{\sum_{j\le M} Y_j,\ M \ge 1\big\}$ and $\big\{W(\sum_{j\le M} T_j),\ M \ge 1\big\}$ are equal in distribution.

Proposition A.8 *For β as in Corollary A.2, Lemma A.3, Lemma A.7 and any $\epsilon > 0$, $\sum_{j\le M_N} T_j - N \ll N^{2\beta+\epsilon}$ a.e.*

Proof Let $\mathcal{A}_M = \sigma\{W(t) : 0 \le t \le \sum_{j\le M} T_j\}$, so $\mathcal{L}_M \subset \mathcal{A}_{\mathcal{M}}$. Each T_j is $\mathcal{A}_j$-measurable, $E(T_j|\mathcal{A}_{j-1}) = E(Y_j^2|\mathcal{L}_{j-1})$ a.e., and $ET_j^p \ll E|Y_j|^{2p}$. In particular, the argument in Lemma A.7 implies that

$$\textstyle\sum_{j\le M}(T_j - E(T_j|\mathcal{A}_{j-1})) \ll N^{2\beta} \text{ a.e.} \qquad \text{(A.2)}$$

Now write

$$\sum T_j - N = \sum(T_j - E(T_j|A_{j-1})) + \sum(E(Y_j^2|\mathcal{L}_{j-1}) - Y_j^2) + \sum Y_j^2 - N.$$

The result follows from (A.2), Corollaries A.2 and A.5, and Lemmas A.3 and A.7. ∎

Theorem A.9 *Let $\beta = \frac{1}{2p} + \frac{1}{4}$ for $2 < p \le 4$ and $\beta = \frac{3}{8}$ for $p > 4$. For any $\epsilon > 0$, $\sum_{n\le N} \eta_n = W(N) + O(N^{\beta+\epsilon})$ a.e.*

Proof By Theorem 2.10 and Corollary A.5, it suffices to prove that $\sum_{j\le M} Y_j = W(N) + O(N^{\beta+2\epsilon})$. Equivalently, $W(\sum_{j\le M} T_j) = W(N) + O(N^{\beta+2\epsilon})$. Hence the result follows from Proposition A.8. ∎

Acknowledgments The research of was supported in part by EPSRC Grant EP/D055520/1 and a Leverhulme Research Fellowship (IM), and by NSF grant DMS-0244529 (MN). IM is grateful to the University of Houston for hospitality during part of this project, and for the use of e-mail given that pine is inadequately supported on the University of Surrey network. MN would like to thank the University of Surrey for hospitality during part of this research.

References

[1] J. Aaronson. *An Introduction to Infinite Ergodic Theory*. Math. Surveys and Monographs **50**, Amer. Math. Soc., 1997.

[2] J. Aaronson and M. Denker. Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps. *Stoch. Dyn.* **1** (2001) 193–237.

[3] M. Benedicks and L.-S. Young. Markov extensions and decay of correlations for certain Hénon maps. *Astérisque* (2000) no. 261, 13–56.

[4] E. Berger. An almost sure invariance principle for stationary ergodic sequences of Banach space valued random variables. *Probab. Theory Related Fields* **84** (1990) 161–201.

[5] I. Berkes and W. Philipp. Approximation theorems for independent and weakly dependent random vectors. *Ann. Probability* **7** (1979) 29–54.

[6] L. A. Bunimovich and Y. G. Sinaĭ. Statistical properties of Lorentz gas with periodic configuration of scatterers. *Comm. Math. Phys.* **78** (1980/81) 479–497.

[7] L. A. Bunimovich, Y. G. Sinaĭ and N. I. Chernov. Statistical properties of two-dimensional hyperbolic billiards. *Uspekhi Mat. Nauk* **46** (1991) 43–92.

[8] D. L. Burkholder. Distribution function inequalities for martingales. *Ann. Probability* **1** (1973) 19–42.

[9] N. I. Chernov. Advanced statistical properties of dispersing billiards. *J. Stat. Phys.* To appear.

[10] N. I. Chernov and D. Dolgopyat. Hyperbolic billiards and statistical physics. Proceedings of International Congress of Mathematicians (Madrid, Spain, 2006).

[11] N. Chernov and L. S. Young. Decay of correlations for Lorentz gases and hard balls. *Hard ball systems and the Lorentz gas*, Encyclopaedia Math. Sci. **101**, Springer, Berlin, 2000, pp. 89–120.

[12] N. I. Chernov and H.-K. Zhang. Billiards with polynomial mixing rates. *Nonlinearity* **18** (2005) 1527–1553.

[13] Y. S. Chow. Local convergence of martingales and the law of large numbers. *Ann. Math. Statist.* **36** (1965) 552–558.

[14] J.-P. Conze and S. Le Borgne. Méthode de martingales et flow géodésique sur une surface de courbure constante négative. *Ergodic Theory Dynam. Systems* **21** (2001) 421–441.

[15] M. Csörgö and P. Révész. A new method to prove Strassen type laws of invariance principle, I. *Z. Wahrscheinlichkeitstheorie verw. Gebiete:* **31** (1975) 255–259.

[16] J. A. Davydov. The invariance principle for stationary processes. *Teor. Verojatnost. i Primenen.* **15** (1970) 498–509.

[17] M. Denker and W. Philipp. Approximation by Brownian motion for Gibbs measures and flows under a function. *Ergodic Theory Dynam. Systems* **4** (1984) 541–552.

[18] D. Dolgopyat. Limit theorems for partially hyperbolic systems. *Trans. Amer. Math. Soc.* **356** (2004) 1637–1689.

[19] A. Dvoretzky. Asymptotic normality for sums of dependent random variables. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory* (Berkeley, Calif.), Univ. California Press, 1972, pp. 513–535.

[20] W. Feller. *An Introduction to Probability Theory and its Applications, II.* Wiley, New York, 1966.

[21] M. J. Field, I. Melbourne and A. Török. Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions. *Ergodic Theory Dynam. Systems* **23** (2003) 87–110.

[22] I. S. Gál and J. F. Koksma. Sur l'ordre de grandeur des fonctions sommables. *Nederl. Akad. Wetensch., Proc.* **53** (1950) 638–653 = Indagationes Math. **12**, 192–207 (1950).

[23] P. Hall and C. C. Heyde. *Martingale Limit Theory and Its Application.* Probability and Mathematical Statistics, Academic Press, New York, 1980.

[24] F. Hofbauer and G. Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. *Math. Z.* **180** (1982) 119–140.

[25] M. Holland and I. Melbourne. Central limit theorems and invariance principles for Lorenz attractors. Preprint, 2006.

[26] N. C. Jain, K. Jogdeo and W. F. Stout. Upper and lower functions for martingales and mixing processes. *Ann. Probability* **3** (1975) 119–145.

[27] J. Kiefer. Skorohod embedding of multivariate rv's, and the sample df. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24** (1972) 1–35.

[28] J. Kuelbs and Walter Philipp. Almost sure invariance principles for partial sums of mixing B-valued random variables. *Ann. Probability* **8** (1980) 1003–1036.

[29] I. Melbourne. Rapid decay of correlations for nonuniformly hyperbolic flows. *Trans. Amer. Math. Soc.* To appear.

[30] I. Melbourne and M. Nicol. Almost sure invariance principle for nonuniformly hyperbolic systems. *Commun. Math. Phys.* **260** (2005) 131–146.

[31] I. Melbourne and A. Török. Central limit theorems and invariance principles for time-one maps of hyperbolic flows. *Commun. Math. Phys.* **229** (2002) 57–71.

[32] I. Melbourne and A. Török. Statistical limit theorems for suspension flows. *Israel J. Math.* **144** (2004) 191–209.

[33] C. A. Morales, M. J. Pacifico and E. R. Pujals. Singular hyperbolic systems. *Proc. Amer. Math. Soc.* **127** (1999) 3393–3401.

[34] G. Morrow and W. Philipp. An almost sure invariance principle for Hilbert space valued martingales. *Trans. Amer. Math. Soc.* **273** (1982) 231–251.

[35] N. Nagayama. Almost sure invariance principle for dynamical systems with stretched exponential mixing rates. *Hiroshima Math. J.* **34** (2004) 371–411.

[36] W. Parry and M. Pollicott. *Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics.* Astérique **187-188**, Société Mathématique de France, Montrouge, 1990.

[37] M. Peligrad, S. Utev and W. B. Wu. A maximal L_p-inequality for stationary sequences and its applications. *Proc. Amer. Math. Soc.* (2006) to appear.

[38] W. Philipp and W. F. Stout. *Almost Sure Invariance Principles for Partial Sums of Weakly Dependent Random Variables.* Memoirs of the Amer. Math. Soc. **161**, Amer. Math. Soc., Providence, RI, 1975.

[39] Y. G. Sinaĭ. Dynamical systems with elastic reflections. Ergodic properties of dispersing billiards. *Uspehi Mat. Nauk* **25** (1970) 141–192.

[40] V. Strassen. An invariance principle for the law of the iterated logarithm. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **3** (1964) 211–226 (1964).

[41] V. Strassen. Almost sure behavior of sums of independent random variables and martingales. *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66)*, Univ. California Press, Berkeley, Calif., 1967, pp. Vol. II: Contributions to Probability Theory, Part 1, pp. 315–343.

[42] D. Szász and T. Varjú. Limit laws and recurrence for the planar Lorentz process with infinite horizon. Preprint.

[43] V. A. Volkonskiĭ and Y. A. Rozanov. Some limit theorems for random functions. I. *Theor. Probability Appl.* **4** (1959) 178–197.

[44] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.* **147** (1998) 585–650.

[45] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.* **110** (1999) 153–188.